Proposed Decision Memo for Heartsbreath Test for Heart Transplant Rejection (CAG-00394N)

Decision Summary

The Centers for Medicare and Medicaid Services (CMS) has reviewed Menssana's request for a national coverage determination (NCD) for the Heartsbreath diagnostic test used as an adjunct to the endomyocardial biopsy to detect grade 3 heart transplant rejection in patients who have had a heart transplant within the last year and an endomyocardial biopsy within the prior month. We believe that the available evidence does not adequately define the technical characteristics of the test nor demonstrate that Heartsbreath testing to predict grade 3 heart transplant rejection improves health outcomes in Medicare beneficiaries. Therefore, we are proposing that the Heartsbreath diagnostic test is not reasonable and necessary under section 1862(a)(1)(A) of the Social Security Act. We are soliciting public comments on this proposed decision pursuant to section 1862(l) of the Social Security Act.

We are particularly interested in receiving additional peer reviewed and published evidence addressing the technical characteristics of the Heartsbreath test during the thirty-day comment period. It is possible that, if additional evidence is provided, Coverage with Evidence Development (CED) using Coverage with Study Participation (CSP) could be considered in the final decision. The current record of evidence, however, does not support coverage under sections 1862(a)(1)(A) and 1862(a)(1)(E) of the Social Security Act. Accordingly, we are requesting the submission of additional information, as well as soliciting public comment on our proposal to non-cover the Heartsbreath diagnostic test.

Back to Top

Proposed Decision Memo

TO: Administrative File: (CAG-00394N)

Heartsbreath Test for Heart Transplant Rejection

FROM: Steve E. Phurrough, MD, MPA

Director, Coverage and Analysis Group

Marcel E. Salive, MD, MPH Director, Division of Medical and Surgical Services

Sandy Jones, RN, MS Lead Analyst

Lori Paserchia, MD Lead Medical Officer

SUBJECT: Proposed Decision Memorandum for Heartsbreath Test for Heart

Transplant Rejection

DATE: September 29, 2008

I. Proposed Decision

The Centers for Medicare and Medicaid Services (CMS) has reviewed Menssana's request for a national coverage determination (NCD) for the Heartsbreath diagnostic test used as an adjunct to the endomyocardial biopsy to detect grade 3 heart transplant rejection in patients who have had a heart transplant within the last year and an endomyocardial biopsy within the prior month. We believe that the available evidence does not adequately define the technical characteristics of the test nor demonstrate that Heartsbreath testing to predict grade 3 heart transplant rejection improves health outcomes in Medicare beneficiaries. Therefore, we are proposing that the Heartsbreath diagnostic test is not reasonable and necessary under section 1862(a)(1)(A) of the Social Security Act. We are soliciting public comments on this proposed decision pursuant to section 1862(I) of the Social Security Act.

Printed on 8/12/2011. Page 3 of 36

We are particularly interested in receiving additional peer reviewed and published evidence addressing the technical characteristics of the Heartsbreath test during the thirty-day comment period. It is possible that, if additional evidence is provided, Coverage with Evidence Development (CED) using Coverage with Study Participation (CSP) could be considered in the final decision. The current record of evidence, however, does not support coverage under sections 1862(a)(1)(A) and 1862(a)(1)(E) of the Social Security Act. Accordingly, we are requesting the submission of additional information, as well as soliciting public comment on our proposal to non-cover the Heartsbreath diagnostic test.

II. Background

The International Society for Heart and Lung Transplantation (ISHLT) reports world-wide there have been over 61,000 heart transplantations completed since the first procedure forty one years ago. In the U.S. there are about 18,000 people living with a heart transplant; and each year, there are an additional 2,000 to 3,000 conducted heart transplants (Edwards 2006). The numbers of Medicare beneficiaries who have received heart transplants over the last three years have varied and include 990 in 2005, 934 in 2006 and 905 in 2007.

A major concern after any transplantation is the rejection of the transplanted organ by the patient's immunologic system. The greatest risk for heart transplant rejection occurs within the first year of transplantation. Since it is difficult to detect rejection using only clinical assessment, endomyocardial biopsy of the right ventricle is typically performed frequently during the first year after transplantation and then periodically afterwards (Phillips, et al. 2004). During this procedure, four to eight small samples of heart tissue are removed (McAllister 1995). Using the ISHLT rating scale,² a pathologist scores the biopsy for the degree of rejection.

Although endomyocardial biopsies are considered by the transplant community to be the gold standard for detecting all grades of heart transplant rejection, the use of biopsy is not without risks and limitations. Biopsy is an invasive procedure that must be performed repeatedly, especially during the first year after transplantation. Some risks of collecting a biopsy include infection, an irregular heartbeat and rupture of the heart wall. Failure to obtain samples during biopsy that contain signs of rejection even when rejection is present is a potential limitation. Another limitation is the inconsistent reading of the same biopsy samples by different pathologists, which is an ongoing concern in the transplant community (Phillips, et al. 2004).

The results of an endomyocardial biopsy are used to decide if the patient is receiving the appropriate amount of immunosuppressive therapy. A change in immunosuppressive therapy appears to be based on the detection of grade 2 or grade 3 heart transplant rejection (Winters, McManus 1996). This reflects a lack of consensus in the transplant community as to what degree of rejection should trigger the need to modify immunosuppressive therapy. It should also be noted that exactly how therapy is modified varies between transplant centers.

The Heartsbreath test is an emerging test to noninvasively detect the presence of heart transplant rejection by collecting breath samples from the patient, which can be performed in virtually any patient care setting. The analysis of the sample is complex and performed in the laboratory with results provided to the clinician shortly thereafter. Goals of the Heartsbreath test, as reported by the requestor and colleagues (2004), are to use the breath test results to detect heart transplant rejection and to decrease the number of invasive endomyocardial biopsies required to monitor for rejection in a patient who has received a heart transplant.

As presented by Phillips and associates (2004) in the HARDBALL study, the Heartsbreath test is based on the understanding that oxidative stress occurs with organ rejection and produces volatile organic compounds (VOCs) that are subsequently eliminated from the body via the breath. These VOCs are identified and measured through technical and statistical analysis to predict heart transplant rejection.

Furthermore, Phillips (2002) describes how to do the Heartsbreath test. The first component of the test requires the collection of two samples. One sample is breath which contains low concentrations of VOCs trapped in a portable apparatus. The other sample is room air that is similarly collected at the same time and contains VOCs. These two samples are compared and an analysis of the 200 plus VOCs is performed using gas chromatography-mass spectroscopy (GC/MS) laboratory techniques in order to determine an assay of the type and quantity of VOCs present in the sample. The result of this laboratory analysis is a set of oxidative stress markers for heart transplant rejection, called the breath methylated alkane contour (BMAC), which is displayed using surface plots and graphical representation of the VOCs. Finally, the Heartsbreath test results were compared to the heart biopsy report to determine the concordance or differences between the findings of the two tests to detect heart transplant rejection. In the HARDBALL study, a BMAC of nine specific VOCs with a mean volume of negative 6.5 was determined to indicate the presence of grade 3A heart rejection in the patients studied. (The evidence Section VII of this PDM further describes these results.)

In a Humanitarian Device Exemption (HDE), the FDA (2004) assessed the Heartsbreath test and approved it for use to assist in the diagnosis of grade 3 heart transplant rejection in patients who have received a heart transplant within the preceding year and an endomyocardial biopsy within the prior month. The Heartsbreath test is intended for use as an adjunct to, and not as a substitute for, endomyocardial biopsy. (Section V of this document provides details about the FDA status of the Heartsbreath test.)

For clinical decision-making using the Heartsbreath test results, the FDA (2004) issued a guide that contains eight possible observation and interpretive outcomes divided into two groups of rejection (less than a grade 3A and grade 3A). Of note, grade 3B or grade 4 rejection were not included because the HARDBALL study, which was the basis for FDA approval, did not investigate patients with more severe grade 3B or grade 4 rejection. The guide is a decision tree (located on-line at: http://www.fda.gov/cdrh/pdf3/H030004c.pdf) that further describes when a second or third biopsy reading must be performed to rule out: 1) an erroneous biopsy reading by a pathologist; 2) a sampling biopsy error; or 3) a Heartsbreath test interpretation error.

The requestor, founder and chief executive officer of Menssana Research, Inc., who is also the lead author of the HARDBALL study, is asking CMS to consider national coverage of the Heartsbreath test as an adjunct to the heart biopsy to detect grade 3 heart transplant rejection in patients who have had a heart transplant within the last year and an endomyocardial biopsy in the prior month.

III. History of Medicare Coverage

There is no national coverage determination (NCD) for the Heartsbreath test and currently there are no local coverage decisions (LCDs) for this emerging technology. On April 10, 2008, CMS accepted this formal request from Menssana Research, Inc. and the first public comment period opened.

Benefit Category

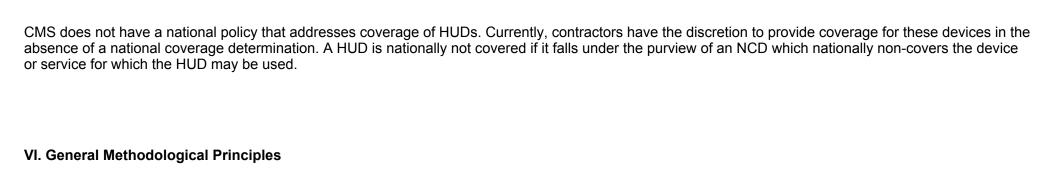
Medicare is a defined benefit program. A prerequisite for Medicare coverage is that an item or service must meet one of the statutorily defined benefit categories in the Social Security Act and not otherwise be excluded from coverage. The Heartsbreath Test at a minimum falls under the benefit category set forth in Title XVIII of the Social Security Act, Section (§)1861(s)(3) (other diagnostic tests), a part B benefit.³

IV. Timeline of Recent Activities

April 10, 2008	CMS accepts Menssana Research, Inc.'s formal NCD request for coverage of the Heartsbreath test for Heart Transplant Rejection. The tracking sheet is posted and the initial 30-day comment period begins.
May 10, 2008	Initial 30 day public comment period closes. Comment is posted on the website.
September 29, 2008	Proposed Decision Memorandum is posted and 30-day public comment period begins.

V. FDA Status

On February 24, 2004, the FDA, Center for Devices and Radiological Health (CDRH), approved Menssana Research Inc.'s HDE application for the Heartsbreath test used to assist in the diagnosis of grade 3 heart transplant rejection in patients who have received heart transplants within the preceding year. The use of the device is limited to patients who have had an endomyocardial biopsy "gold standard" within the previous month and the Heartsbreath test is intended for use as an adjunct to, and not as a substitute for, endomyocardial biopsy. The FDA letter⁴ refers to this application submitted by Menssana Research, Inc. and the public was notified of this FDA decision.⁵ Based on the data submitted with the HDE application, the FDA determined that "the Heartsbreath test for heart transplant rejection will not expose patients to an unreasonable or significant risk of illness or injury and that, when used following the instructions for use, that there is a probable benefit to health that outweighs the risks of illness or injury."⁶



When making national coverage decisions, CMS generally evaluates relevant clinical evidence to determine whether or not the evidence is of sufficient quality to support a finding that an item or service falling within a benefit category is reasonable and necessary for the diagnosis or treatment of an illness or injury or to improve the functioning of a malformed body member. The evidence may consist of external technology assessments, internal review of published and unpublished studies, recommendations from the Medicare Coverage Advisory Committee, evidence-based guidelines, professional society position statements, expert opinion and public comments. The critical appraisal of the evidence is to determine to what degree we are confident that: 1) the specific clinical questions relevant to the coverage request can be answered conclusively; and 2) the intervention will improve patients' health outcomes. (The General Methodological Principles of Study Design is located in *Appendix A*.)

We divide the assessment of clinical evidence into three stages: 1) the quality of the individual studies; 2) the relevance of findings from individual studies to the Medicare population; and 3) overarching conclusions that can be drawn from the body of the evidence on the direction and magnitude of the intervention's risks and benefits.

Public comments sometimes cite the published clinical evidence and gives CMS useful information. Public comments that give information on unpublished evidence such as results of individual practitioners or patients are less rigorous and therefore less useful when making a coverage determination. CMS uses the initial public comments to inform its proposed decision. CMS responds in detail to the public comments on a proposed decision when issuing the final decision memorandum.

VII. Evidence

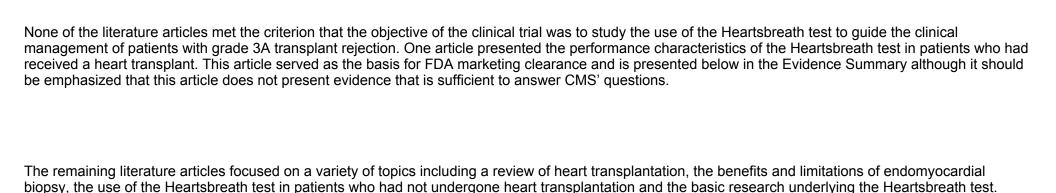
A. Introduction This analysis focuses on whether the Heartsbreath test is useful to guide clinical decision-making and patient management within one year of heart transplantation. In evaluating diagnostic tests, Mol and colleagues (2003) reported: "Whether or not patients are better off from undergoing a diagnostic test will depend on how test information is used to guide subsequent decisions on starting, stopping or modifying treatment. Consequently, the practical value of a diagnostic test can only be assessed by taking into account subsequent health outcomes." When a proven, well established association or pathway is available, intermediate health outcomes may also be considered. For example, if a particular diagnostic test result can be shown to change patient management and other evidence has demonstrated that those patient management changes improve health outcomes, then those separate sources of evidence may be sufficient to demonstrate positive health outcomes from the diagnostic test. Literature Search On June 16, 2008, CMS performed a PubMed search of the literature using the following search terms: "breath test" and "heart transplant" or "heart transplant" rejection." The limitations used were: Human, English, and Article type (Clinical Trial, Randomized Clinical Trial, Meta-analysis, Review, Practice Guideline). B. Discussion of evidence reviewed

• Is the evidence adequate to conclude that heart transplant patients whose post transplant testing management includes Heartsbreath testing experience improved health outcomes compared to patients whose management does not include Heartsbreath testing?

Printed on 8/12/2011. Page 9 of 36

1. Questions

0	Does a negative Heartsbreath test sufficiently exclude grade 3A rejection so as to obviate the need for endomyocardial biopsy in a patient who would otherwise be biopsied? Does a positive Heartsbreath test sufficiently diagnose grade 3A rejection so as to inform immunosuppressive therapy without the need for an endomyocardial biopsy?
2. External tec	hnology assessment
We did not requ	uest an external technology assessment on this issue and are not aware of any other similar assessments.
On July 14, 200 assessments w	08, the Cochrane online database and the NICE online database were searched using the term "cardiac allograft rejection." No technology rere found.
3. Internal tech	nnology assessment
The internal tec	chnology assessment was based on the results of the CMS literature search and literature articles submitted by the requestor.
	e article sources, CMS looked for published, peer-reviewed evidence of controlled clinical trials that provided results on the use of the est to guide the clinical management of patients with grade 3A heart transplant rejection.



Evidence Summary

Phillips M, et al. Heart allograft rejection: Detection with breath alkanes in low levels (the HARDBALL study). International Society for Heart and Lung Transplantation 2004;23:701-8.

This was a prospective, multi-site study of 539 heart transplant recipients with the goal to determine if the BMAC could be used in patients as a marker of heart transplant rejection. An age-matched group of 32 healthy volunteers also received the Heartsbreath test. A total of 1061 technically satisfactory Heartsbreath (BMAC) samples were collected prior to endomyocardial biopsy.

The endomyocardial biopsies were graded for severity of rejection by each study site's pathologist using the Billingham version of the ISHLT classification. The site pathologists then selected and forwarded the most representative biopsy slide to two offsite but centrally-located cardiac pathologists (i.e., the expert pathologists). Each cardiac pathologist performed an independent evaluation using the ISHLT classification; discordant interpretations were reviewed jointly until a concordant interpretation was obtained. The cardiac pathologists provided the expert reading of the biopsy and their concordant reading was considered to be the "gold standard" interpretation of the severity of rejection. The site pathologists were blinded to the results of the Heartsbreath test while the expert pathologists were completely blinded to all patient information.

Analysis of the BMAC samples was conducted by two investigators who were blinded to the results of the endomyocardial biopsies. BMAC samples were analyzed by type of group: BMACs from patients with grade 3A rejection; BMACs from patients with grade 0, 1 or 2 rejection; BMACs from healthy volunteers. The BMAC samples from the patients with grade 3A rejections were compared to those samples from patients with grade 0, 1 or 2 rejection using forward stepwise discriminant analysis. The resultant model reported a value ranging from zero to one that indicated the probability of grade 3 rejection for each BMAC. A cross-validation was also performed using another type of discriminant analysis. The authors constructed three-dimensional volume under curve surface plots of the BMAC data for each of the three groups to display the polarity (direction) and distribution (dispersion) of the VOCs.

The results from the Heartsbreath test (BMAC samples) were compared to the gold standard endomyocardial biopsy results and performance characteristics (sensitivity, specificity, negative predictive value, positive predictive value) were determined for each. The results of the site pathologists' reading of the endomyocardial biopsy were also compared to the gold standard biopsy results and the types of performance characteristics were calculated.

For the 539 heart transplant patients, the mean age was 54 years and 128 were women. For the 32 healthy volunteers, the mean age was 53 years and sixteen were women. Four percent of the 1061 biopsy samples demonstrated grade 3A rejection. Of the remaining samples, 60.8% had grade 0, 26.5% had grade 1 and 8.8% had grade 2 rejection. No patient with grade 3B or 4 rejection was found. There was no significant difference in the mean age between those patients with grade 3A rejection and those with grade 0, 1 or 2 rejection.

The performance characteristics of the Heartsbreath test and the site pathologist were determined to be:

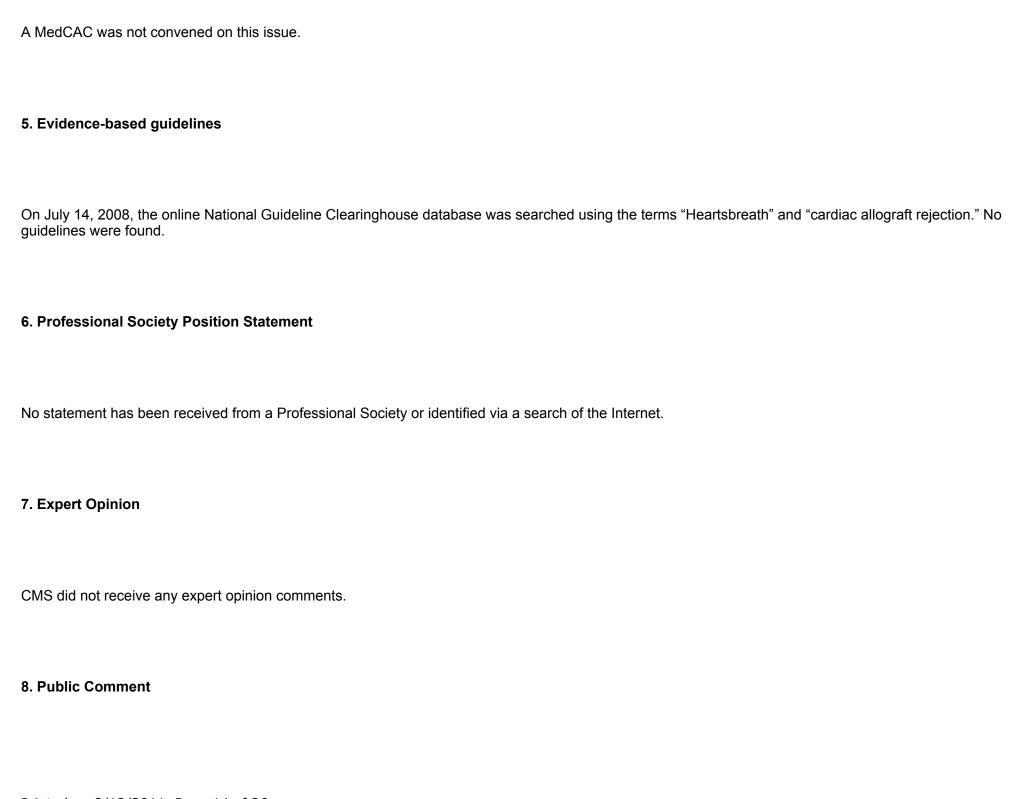
	Site Pathologists	Heartsbreath test
Performance Characteristic		
	42.4%	59.5%
Sensitivity		

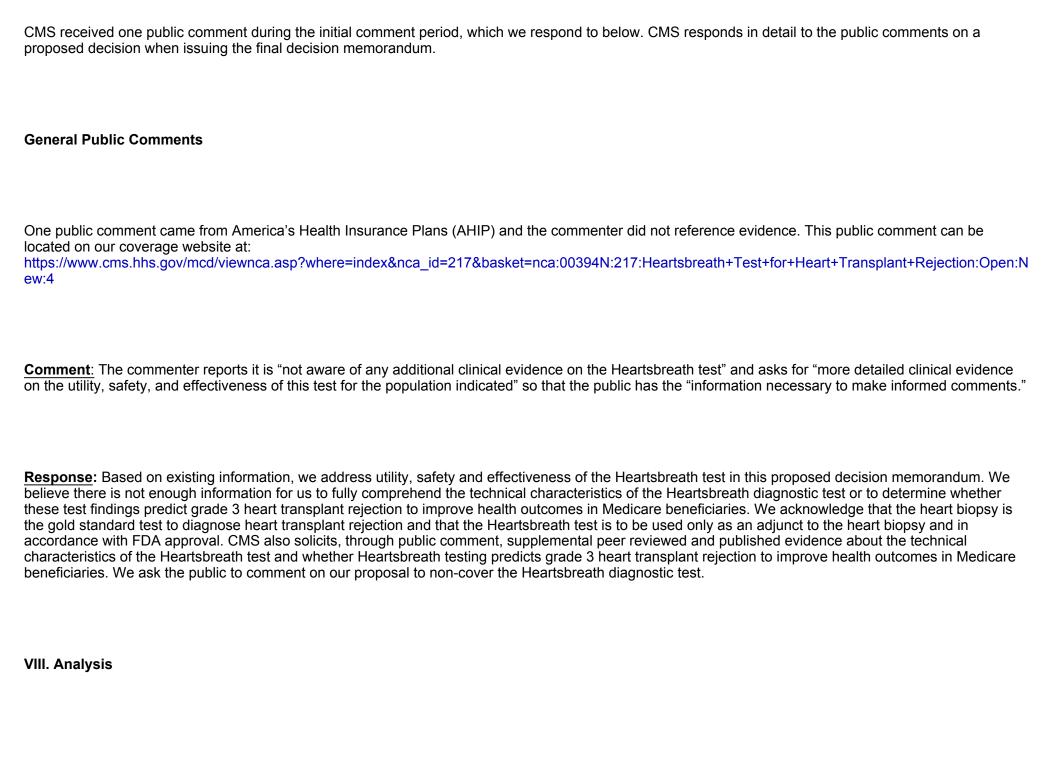
Specificity	97.0%	58.8%
Positive Predictive Value	45.2%	5.6%
Negative Predictive Value	96.7%	97.2%

Nine VOCs in the BMAC samples were identified to be associated with grade 3A rejection. The surface plot and the mean volume for the BMACs from patients with grade 3A rejection showed a significantly different pattern in terms of polarity (referred to by the authors as a paradoxical reversal) and distribution compared to the surface plot and mean volume for the BMACs from patients with grade 0, 1 or 2 rejection. The surface plot and mean volume for the BMACs from the healthy volunteers was similar in terms of polarity to that for the patients with grade 3A rejection although the magnitude of the mean volume and the distribution were different.

The authors stated that in the clinical setting a positive Heartsbreath test should be followed by an endomyocardial biopsy because this would increase the positive predictive value (probability of finding grade 3A rejection) from 5.6% to 45.2%. Alternatively, a negative Heartsbreath test, with a 97.2% negative predictive value (the degree of confidence that no grade 3A rejection is present), would obviate the need for a subsequent biopsy. The authors also postulated that a change in the ability of a patient with grade 3A rejection to metabolize VOCs may be the reason for the unexpected difference in the surface plot and mean volume between the BMACs from patients with grade 3A rejection and those with grade 0, 1 or 2 rejection. The authors concluded that the Heartsbreath test "could potentially identify transplant recipients at low risk of grade 3A rejection and reduce the number of endomyocardial biopsies."

4. Medicare Evidence Development and Coverage Advisory Committee (MedCAC) Meeting





National coverage determinations (NCDs) are determinations by the Secretary with respect to whether or not a particular item or service is covered nationally by Medicare (§1869(f)(1)(B) of the Act). In order to be covered by Medicare, an item or service must fall within one or more benefit categories contained within Part A or Part B, and must not be otherwise excluded from coverage. Moreover, with limited exceptions, the expenses incurred for items or services must be "reasonable and necessary for the diagnosis or treatment of illness or injury or to improve the functioning of a malformed body member." See §1862(a)(1)(A) of the Act. This section presents the agency's evaluation of the evidence considered and conclusions reached for the assessment questions.

As a diagnostic test, the Heartsbreath test would not be expected to directly change health outcomes. Rather, a diagnostic test affects health outcomes through changes in disease management brought about by physician actions taken in response to test results. Such actions may include decisions to treat or withhold treatment, to choose one treatment modality over another, or to choose a different dose or duration of the same treatment. To some extent the usefulness of a test result is constrained by the available treatment options. As noted in the Background section, the number of practical treatment options for transplant rejection is limited. A patient whose rejection is not readily controlled with a particular regimen is likely to be prescribed alternative or additional drug treatment. In addressing the question, one of the factors we consider is whether there is sufficient evidence that the incremental information derived from Heartsbreath testing leads to improved control of transplant rejection by causing physicians to prescribe a different anti-rejection regimen than they would have prescribed without access to Heartsbreath test results, or to forego invasive endomyocardial biopsy.

The Medicare regulations at 42 CFR 410.32(a) state in part, "...diagnostic tests must be ordered by the physician who is treating the beneficiary, that is, the physician who furnishes a consultation or treats a beneficiary for a specific medical problem and who uses the results in the management of the beneficiary's specific medical problem." Thus we look for evidence demonstrating how the treating physician uses the result of a Heartsbreath test to manage the anti-rejection treatment in patients who have undergone heart transplant.

Ideally we would see evidence that the systematic incorporation of Heartsbreath test results into an anti-rejection treatment algorithm leads treating physicians to prescribe different classes of medications or more appropriate dosages of the same medications than they would otherwise have prescribed, and that patients whose treatment is changed by Heartsbreath test results remain on the regimen and achieve better long term anti-rejection control documented by repeated assessments over time. Unfortunately the data are silent on health outcomes, and do not establish that the treating physicians currently base patient management on the Heartsbreath test result.

We considered the evidence in the hierarchical framework of Fryback and Thornbury (1991) where Level 2 evidence addresses diagnostic accuracy, sensitivity, and specificity of the test; Level 3 evidence focuses on whether the information produces change in the physician's diagnostic thinking; Level 4 evidence concerns the effect on the patient management plan and Level 5 evidence measures the effect of the diagnostic information on patient outcomes. Most studies have focused on test characteristics and have not considered health outcomes, such as mortality, morbidity or reduction of invasive biopsy. We believe that health outcomes is more persuasive than test characteristics.

CMS asked the following questions when analyzing the evidence:

- Is the evidence adequate to conclude that heart transplant patients whose post transplant testing management includes Heartsbreath testing experience improved health outcomes compared to patients whose management does not include Heartsbreath testing?
 - Does a negative Heartsbreath test sufficiently exclude grade 3A rejection so as to obviate the need for endomyocardial biopsy in a patient who would otherwise be biopsied?
 - Does a positive Heartsbreath test sufficiently diagnose grade 3A rejection so as to inform immunosuppressive therapy without the need for an endomyocardial biopsy?

CMS found one study, the HARDBALL study, which examined the Heartsbreath test in patients who had received a heart transplant. The study reported the technical characteristics of the Heartsbreath test as well as its performance characteristics compared to endomyocardial biopsy. CMS has three major concerns about the HARDBALL study. Two concerns relate to the reported results of the study and one relates to the study design.

The first major concern CMS has about the HARDBALL study centers on the technical characteristics of the Heartsbreath test. By technical characteristics CMS is referring to the chemical/biochemical and physical aspects of a test method. For the Heartsbreath test, the chemical/biochemical and physical aspects include the collection of the patient's breath and the analysis of the sample.

During the HARDBALL study, statistical analysis was employed to determine which of the many VOCs typically present in a breath sample were most associated with the presence of grade 3A rejection in patients with a heart transplant. Nine specific VOCs were identified. It is unclear if these specific VOCs are representative of grade 3A rejection in all patients with a heart transplant or are representative of only the sample of patients in the HARDBALL study. In other words, it remains unknown if this result from the study can, or should, be generalized and applied to future patients. Therefore, the specific number and types of VOCs found to be representative of grade 3A rejection should be confirmed in a subsequent study.

In addition, the finding of paradoxical reversal of the BMACs from patients with grade 3A rejection compared to patients with less than grade 3A rejection, a result that the authors noted was unexpected, calls into question if the technical characteristics of the Heartsbreath test are sufficiently understood for this specific patient population. The authors postulated that an increase in the metabolism of the VOCs may have produced the paradoxical reversal and cited a well-known example of such a metabolic change leading to decreased blood concentrations of a typical immunosuppressive drug to support their hypothesis. However, this is just a hypothesis and should be investigated by conducting more proof-of-concept studies.

Consequently, CMS asks the public if additional evidence in support of the technical characteristics in this specific patient population is available but was not apparent to CMS during its review. If not, then additional proof-of-concept investigations are warranted and should be conducted prior to proceeding with a clinical trial that studies whether the Heartsbreath test improves health outcomes for patients with heart transplant when used as an adjunct or as a substitute for biopsy.

The second major concern CMS has about the HARDBALL study centers on the report of its performance or test characteristics. Endomyocardial biopsy and the Heartsbreath test are diagnostic tests. The objective of a diagnostic test is to accurately find disease when it is present in the patient and exclude it when it is absent. The probability of a positive test in a patient with the disease is called the sensitivity of the test. The probability of a negative test in a patient without the disease is called the specificity of the test. Hence, sensitivity and specificity are two measures of the validity of a diagnostic test (Hennekens, et al. 1987). In a related fashion, the probability of disease being present in a patient when the test is positive is called the positive predictive value. The probability of no disease being present in a patient when the test is negative predictive value. In this PDM, these four attributes of a test are referred to as the performance or test characteristics. The positive and negative predictive values of a test will vary with the prevalence of the disease in the population being tested (Hennekens, et al. 1987).

For a diagnostic test, a high sensitivity decreases the chance that a diseased patient will have a negative test. This is particularly desirable when the disease itself is potentially lethal yet can be treated safely and effectively. This is a critical point for patients who have had a heart transplant where it is imperative to detect high grade (e.g., grade 3) rejection when it is present because it can have serious consequences for the viability of the transplant and the patient.

Using the cardiac pathologist read biopsy result as the standard, the sensitivity was only 59.5% for the Heartsbreath test in the HARDBALL study. Interestingly, the sensitivity was originally 78.6% but was decreased after cross-validation. The authors did not provide an explanation for the significant decrease in the sensitivity.

As the authors of the HARDBALL study noted, endomyocardial biopsy is still considered to be the gold standard for the diagnosis of rejection hence there is no apparent consensus for an alternative comparator. Using the cardiac pathologist read biopsy result as the standard, the sensitivity of endomyocardial biopsy as read by the site pathologists was even lower (42.4%). The significant lack of concordance between the site pathologists and the expert pathologists exemplifies the controversy in the transplant community regarding the limitations of endomyocardial biopsy. It also points out the apparently unmet need for a better gold standard comparator than endomyocardial biopsy.

The authors of HARDBALL propose to use the Heartsbreath test as an adjunct to endomyocardial biopsy but only when the Heartsbreath test is positive, i.e., Heartsbreath negative patients would not be biopsied. Their rationale stems from the high negative predictive value (NPV) (> 95%) found for the Heartsbreath test. The high NPV is not surprising given that only four percent of the study population had grade 3A rejection. From a statistical point-of-view, this very low prevalence of rejection has a large negative impact on the positive predictive value of the test.

The low positive predictive value indicates that the post-test probability that a patient has grade 3A rejection remains very low even when the Heartsbreath test is positive in the population tested. The probability of rejection is relatively higher with a positive biopsy, though still low in absolute terms. Therefore, the strategy is to acknowledge that the Heartsbreath test is inadequate to detect grade 3A rejection when the patient actually does have grade 3A rejection and therefore to perform an endomyocardial biopsy, which has a relatively better ability to detect grade 3A rejection when the patient actually does have grade 3A rejection.

Interestingly, the FDA labeling for the Heartsbreath test states that it should only be used as an adjunct to endomyocardial biopsy and only after biopsy is performed. If biopsy is indeed the gold standard it is unclear if or how the Heartsbreath test result would be used by the treating physician in the management of the patient.

CMS believes that the putative clinical value of the Heartsbreath test would be to differentiate transplant patients who have grade 3A rejection from those who do not with sufficient accuracy so that the treating physician could confidently base treatment decisions on the results of the Heartsbreath test without the need to also perform an endomyocardial biopsy. CMS acknowledges that although biopsy has its own limitations, it is nonetheless the current gold standard for diagnosing rejection.

It is important to note that no patients in the HARDBALL study had grade 3B or grade 4 rejection. Hence, the Heartsbreath test is currently designed to predict if a patient has grade 3A rejection or not. The clinical utility of this narrow determination is unknown, especially given the controversy in the transplant community regarding what grade of rejection should indicate a change in clinical management.

Printed on 8/12/2011. Page 19 of 36

Lastly, the HARDBALL study used the Billingham ISHLT classification rating scale to assign the severity of rejection. The impact of the subsequent revisions to this ISHLT rejection classification rating scale (footnote 2 of this PDM) on the Heartsbreath test performance characteristics and, more importantly, on the application of Heartsbreath test results to patient management is unknown and needs to be explored.

CMS' third major concern with the HARDBALL study centers on the study design. While the study produced results that suggest how the Heartsbreath test can be used as an adjunct or substitute to endomyocardial biopsy, the design of this study did not allow for a determination of whether using the Heartsbreath test as an adjunct or substitute would lead to an improved health outcome. In other words, it is still unknown if the Heartsbreath test, when used to guide clinical management of a patient after heart transplant is beneficial in reducing morbidity or mortality. This type of evidence is an important factor to CMS when determining if a diagnostic test is reasonable and necessary.

For these reasons, the evidence is insufficient to conclude that heart transplant patients whose post transplant testing management includes Heartsbreath testing experience improved health outcomes compared to patients whose management does not include Heartsbreath testing. A negative Heartsbreath test does not sufficiently exclude Grade 3A rejection so as to obviate the need for endomyocardial biopsy in a patient who would otherwise be biopsied. A positive Heartsbreath test does not sufficiently diagnose Grade 3A rejection so as to inform immunosuppressive therapy without the need for an endomyocardial biopsy. Thus, the evidence is inadequate to conclude that the Heartsbreath test is reasonable and necessary under section 1862(a)(1)(A) for the diagnosis of heart transplant rejection.

While adequate health outcomes have not been studied, the clinical results of the HARDBALL study have shown promise for a technology that may have potential as a noninvasive approach for the diagnosis of grade 3A heart transplant rejection. A noninvasive test is preferable if it can yield the same/similar benefit as an invasive test. In this case, the potential complications of endomyocardial biopsy such as heart perforation, bleeding and infection can be avoided.

However, the promise of benefit is premised on clear technical characteristics in this patient population. Our review of the evidence leaves us concerned that too little is known about this test's ability to accurately diagnose rejection in heart transplant patients. Therefore, during the thirty-day comment period for the proposed decision memorandum, CMS is interested in receiving additional peer-reviewed and published evidence that addresses our concerns about the technical characteristics of the Heartsbreath test. With the submission of additional evidence during this time period that satisfies our concerns about the technical characteristics, it is possible that CMS may consider Coverage with Evidence Development using Coverage with Study Participation in the final decision memorandum. If further evidence is not available, then additional proof-of-concept investigations are warranted and should be conducted prior to proceeding with a clinical trial to study whether the Heartsbreath test improves health outcomes for patients with heart transplant when used as an adjunct or as a substitute for biopsy.

IX. Summary

The Centers for Medicare and Medicaid Services (CMS) has reviewed Menssana's request for a national coverage determination (NCD) for the Heartsbreath diagnostic test used as an adjunct to the endomyocardial biopsy to detect grade 3 heart transplant rejection in patients who have had a heart transplant within the last year and an endomyocardial biopsy within the prior month. We believe that the available evidence does not adequately define the technical characteristics of the test nor demonstrate that Heartsbreath testing to predict grade 3 heart transplant rejection improves health outcomes in Medicare beneficiaries. Therefore, we are proposing that the Heartsbreath diagnostic test is not reasonable and necessary under section 1862(a)(1)(A) of the Social Security Act. We are soliciting public comments on this proposed decision pursuant to section 1862(I) of the Social Security Act.

We are particularly interested in receiving additional peer reviewed and published evidence addressing the technical characteristics of the Heartsbreath test during the thirty-day comment period. It is possible that, if additional evidence is provided, Coverage with Evidence Development (CED) using Coverage with Study Participation (CSP) could be considered in the final decision. The current record of evidence, however, does not support coverage under sections 1862(a)(1)(A) and 1862(a)(1)(E) of the Social Security Act. Accordingly, we are requesting the submission of additional information, as well as soliciting public comment on our proposal to non-cover the Heartsbreath diagnostic test.

APPENDIX A

General Methodological Principles of Study Design

(Section VI of the Proposed Decision Memorandum)

When making national coverage determinations, CMS evaluates relevant clinical evidence to determine whether or not the evidence is of sufficient quality to support a finding that an item or service falling within a benefit category is reasonable and necessary for the diagnosis or treatment of an illness or injury or to improve the functioning of a malformed body member. The overall objective for the critical appraisal of the evidence is to determine to what degree we are confident that: 1) the specific assessment questions can be answered conclusively; and 2) the intervention will improve health outcomes for patients.

We divide the assessment of clinical evidence into three stages: 1) the quality of the individual studies; 2) the generalizability of findings from individual studies to the Medicare population; and 3) overarching conclusions that can be drawn from the body of the evidence on the direction and magnitude of the intervention's potential risks and benefits.

The methodological principles described below represent a broad discussion of the issues we consider when reviewing clinical evidence. However, it should be noted that each coverage determination has its unique methodological aspects.

Assessing Individual Studies

Methodologists have developed criteria to determine weaknesses and strengths of clinical research. Strength of evidence generally refers to: 1) the scientific validity underlying study findings regarding causal relationships between health care interventions and health outcomes; and 2) the reduction of bias. In general, some of the methodological attributes associated with stronger evidence include those listed below:

- Use of randomization (allocation of patients to either intervention or control group) in order to minimize bias.
- Use of contemporaneous control groups (rather than historical controls) in order to ensure comparability between the intervention and control groups.
- Prospective (rather than retrospective) studies to ensure a more thorough and systematical assessment of factors related to outcomes.
- Larger sample sizes in studies to help ensure adequate numbers of patients are enrolled to demonstrate both statistically significant as well as clinically significant outcomes that can be extrapolated to the Medicare population. Sample size should be large enough to make chance an unlikely explanation for what was found.
- Masking (blinding) to ensure patients and investigators do not know to which group patients were assigned (intervention or control). This is important
 especially in subjective outcomes, such as pain or quality of life, where enthusiasm and psychological factors may lead to an improved perceived
 outcome by either the patient or assessor.

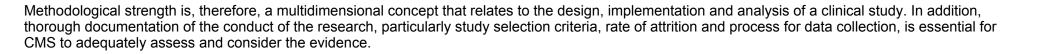
Regardless of whether the design of a study is a randomized controlled trial, a non-randomized controlled trial, a cohort study or a case-control study, the primary criterion for methodological strength or quality is the extent to which differences between intervention and control groups can be attributed to the intervention studied. This is known as internal validity. Various types of bias can undermine internal validity. These include:

- Different characteristics between patients participating and those theoretically eligible for study but not participating (selection bias).
- Co-interventions or provision of care apart from the intervention under evaluation (performance bias).
- Differential assessment of outcome (detection bias).
- Occurrence and reporting of patients who do not complete the study (attrition bias).

In principle, rankings of research design have been based on the ability of each study design category to minimize these biases. A randomized controlled trial minimizes systematic bias (in theory) by selecting a sample of participants from a particular population and allocating them randomly to the intervention and control groups. Thus, in general, randomized controlled studies have been typically assigned the greatest strength, followed by non-randomized clinical trials and controlled observational studies. The design, conduct and analysis of trials are important factors as well. For example, a well designed and conducted observational study with a large sample size may provide stronger evidence than a poorly designed and conducted randomized controlled trial with a small sample size. The following is a representative list of study designs (some of which have alternative names) ranked from most to least methodologically rigorous in their potential ability to minimize systematic bias:

- Randomized controlled trials
- Non-randomized controlled trials
- Prospective cohort studies
- Retrospective case control studies
- Cross-sectional studies
- Surveillance studies (e.g., using registries or surveys)
- Consecutive case series
- Single case reports

When there are merely associations but not causal relationships between a study's variables and outcomes, it is important not to draw causal inferences. Confounding refers to independent variables that systematically vary with the causal variable. This distorts measurement of the outcome of interest because its effect size is mixed with the effects of other extraneous factors. For observational, and in some cases randomized controlled trials, the method in which confounding factors are handled (either through stratification or appropriate statistical modeling) are of particular concern. For example, in order to interpret and generalize conclusions to our population of Medicare patients, it may be necessary for studies to match or stratify their intervention and control groups by patient age or co-morbidities.



Generalizability of Clinical Evidence to the Medicare Population

The applicability of the results of a study to other populations, settings, treatment regimens and outcomes assessed is known as external validity. Even well-designed and well-conducted trials may not supply the evidence needed if the results of a study are not applicable to the Medicare population. Evidence that provides accurate information about a population or setting not well represented in the Medicare program would be considered but would suffer from limited generalizability.

The extent to which the results of a trial are applicable to other circumstances is often a matter of judgment that depends on specific study characteristics, primarily the patient population studied (age, sex, severity of disease and presence of co-morbidities) and the care setting (primary to tertiary level of care, as well as the experience and specialization of the care provider). Additional relevant variables are treatment regimens (dosage, timing and route of administration), co-interventions or concomitant therapies, and type of outcome and length of follow-up.

The level of care and the experience of the providers in the study are other crucial elements in assessing a study's external validity. Trial participants in an academic medical center may receive more or different attention than is typically available in non-tertiary settings. For example, an investigator's lengthy and detailed explanations of the potential benefits of the intervention and/or the use of new equipment provided to the academic center by the study sponsor may raise doubts about the applicability of study findings to community practice.

Given the evidence available in the research literature, some degree of generalization about an intervention's potential benefits and harms is invariably required in making coverage determinations for the Medicare population. Conditions that assist us in making reasonable generalizations are biologic plausibility, similarities between the populations studied and Medicare patients (age, sex, ethnicity and clinical presentation) and similarities of the intervention studied to those that would be routinely available in community practice.

A study's selected outcomes are an important consideration in generalizing available clinical evidence to Medicare coverage determinations. One of the goals of our determination process is to assess health outcomes. We are interested in the results of changed patient management not just altered management. These outcomes include resultant risks and benefits such as increased or decreased morbidity and mortality. In order to make this determination, it is often necessary to evaluate whether the strength of the evidence is adequate to draw conclusions about the direction and magnitude of each individual outcome relevant to the intervention under study. In addition, it is important that an intervention's benefits are clinically significant and durable, rather than marginal or short-lived. Generally, an intervention is not reasonable and necessary if its risks outweigh its benefits.

If key health outcomes have not been studied or the direction of clinical effect is inconclusive, we may also evaluate the strength and adequacy of indirect evidence linking intermediate or surrogate outcomes to our outcomes of interest.

Assessing the Relative Magnitude of Risks and Benefits

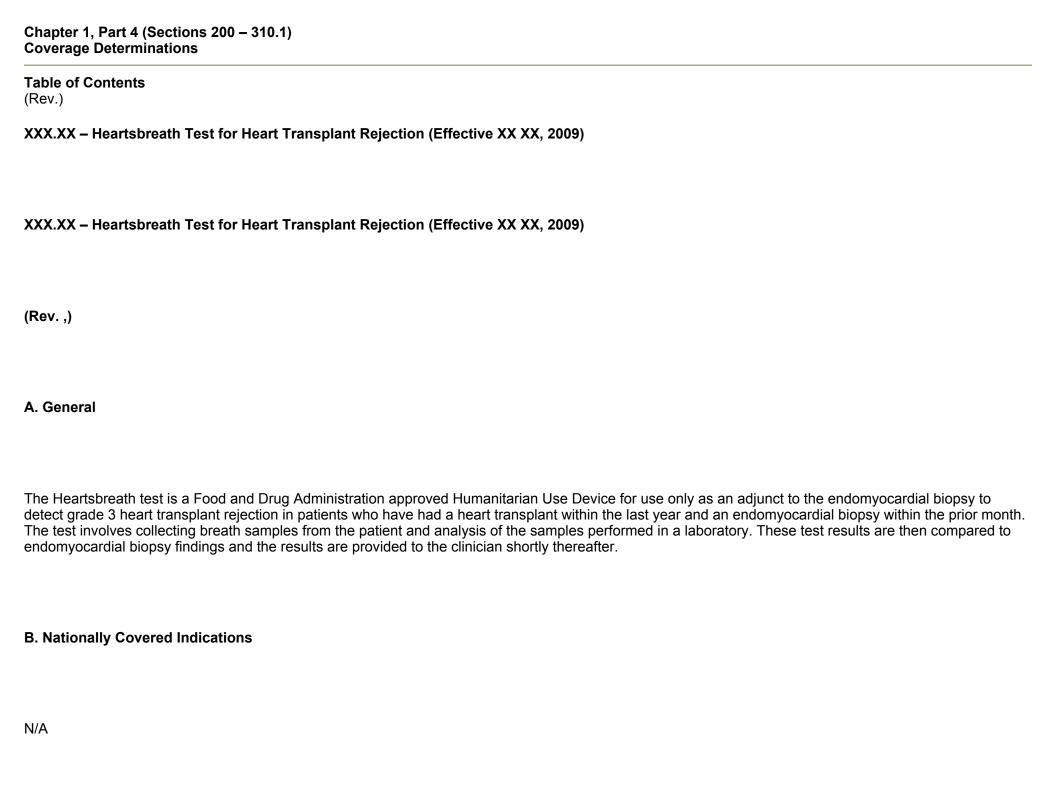
Generally, an intervention is not reasonable and necessary if its risks outweigh its benefits. Health outcomes are one of several considerations in determining whether an item or service is reasonable and necessary. For most determinations, CMS evaluates whether reported benefits translate into improved health outcomes. CMS places greater emphasis on health outcomes actually experienced by patients, such as quality of life, functional status, duration of disability, morbidity and mortality, and less emphasis on outcomes that patients do not directly experience, such as intermediate outcomes, surrogate outcomes, and laboratory or radiographic responses. The direction, magnitude and consistency of the risks and benefits across studies are also important considerations. Based on the analysis of the strength of the evidence, CMS assesses the relative magnitude of an intervention or technology's benefits and risk of harm to Medicare beneficiaries.

Appendix B

DRAFT

Medicare National Coverage Determinations Manual

Printed on 8/12/2011. Page 25 of 36

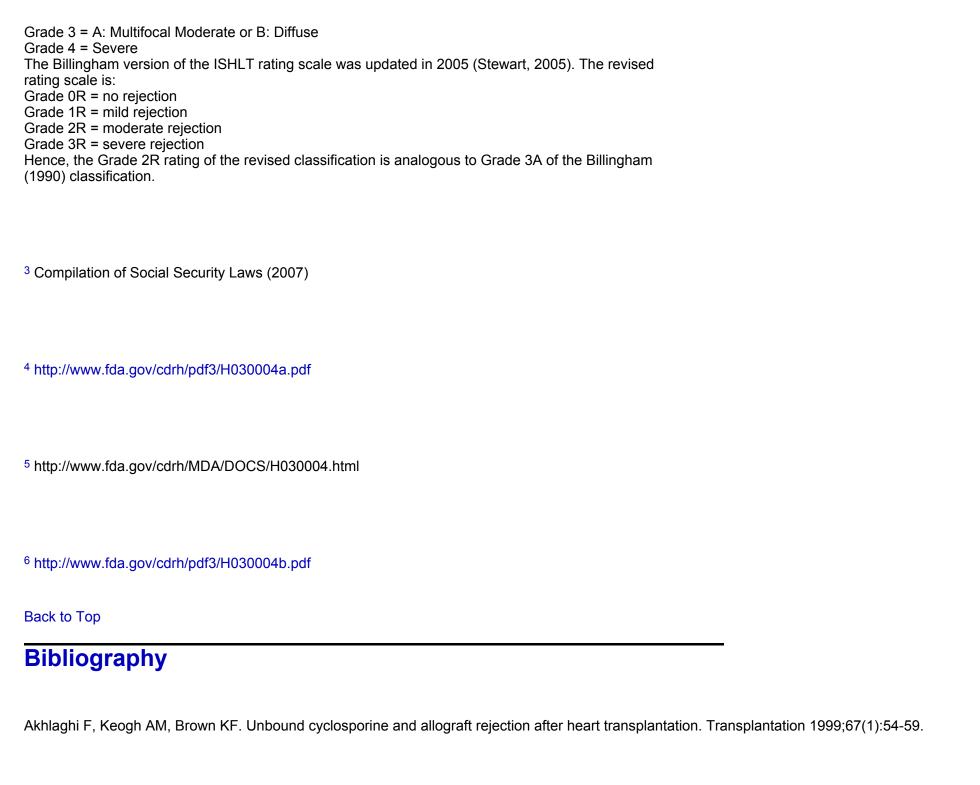


C. Nationally Non-Covered Indications
Effective for services performed on or after January XX, 2009, CMS has determined that the available evidence does not adequately define the technical characteristics of the Heartsbreath test nor demonstrate that Heartsbreath testing to predict grade 3 heart transplant rejection improves health outcomes in Medicare beneficiaries. We conclude that the Heartsbreath test to detect grade 3 heart transplant rejection in patients who have had a heart transplant within the last year and an endomyocardial biopsy within the prior month is not reasonable and necessary under section 1862(a)(1)(A) and the evidence does not support coverage under section 1862(a)(1)(E) of the Social Security Act. Therefore, the Heartsbreath test is noncovered .
D. Other
N/A
(This NCD last reviewed XXXX)
¹ Medicare Part B Systems Extract and Summary System (BESS) Procedure Summary File for Heart Transplants in 2005, 2006 and 2007. Data pulled on September 15, 2008.
² International Society for Heart and Lung Transplantation (ISHLT) rating scale (Billingham 1990)

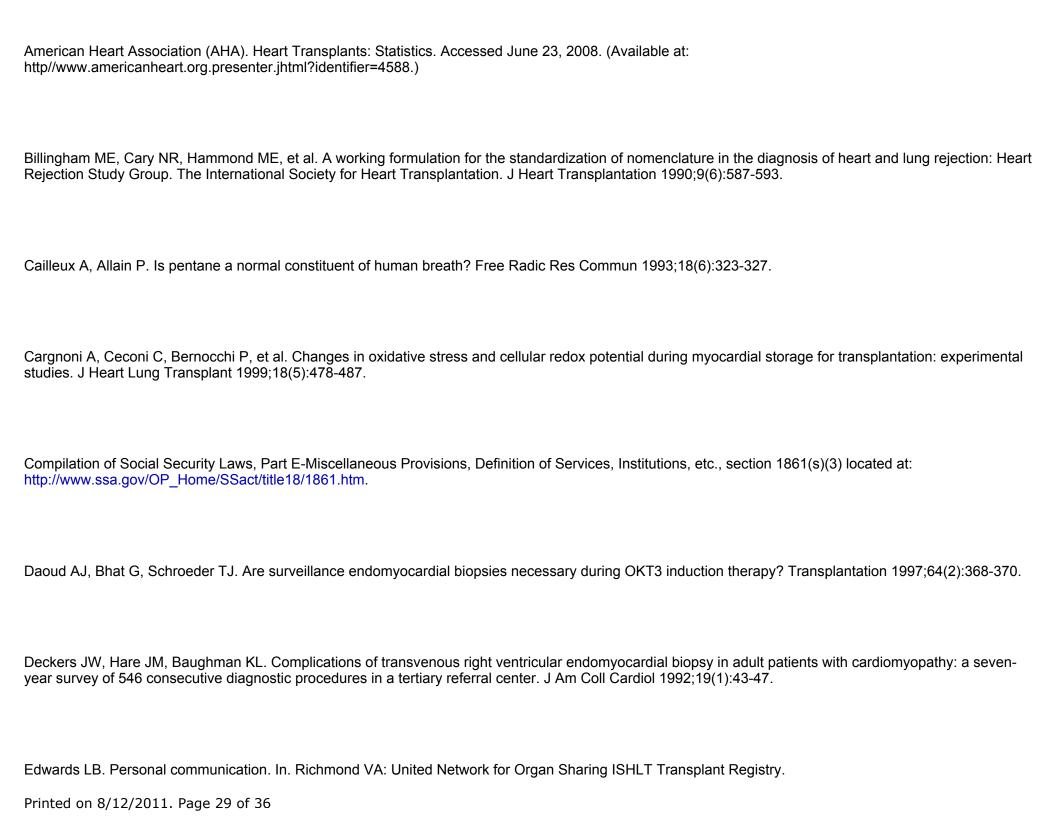
Printed on 8/12/2011. Page 27 of 36

Grade 0 = Absent

Grades 1A, 1B = Mild Grade 2 = Focal Moderate



Printed on 8/12/2011. Page 28 of 36

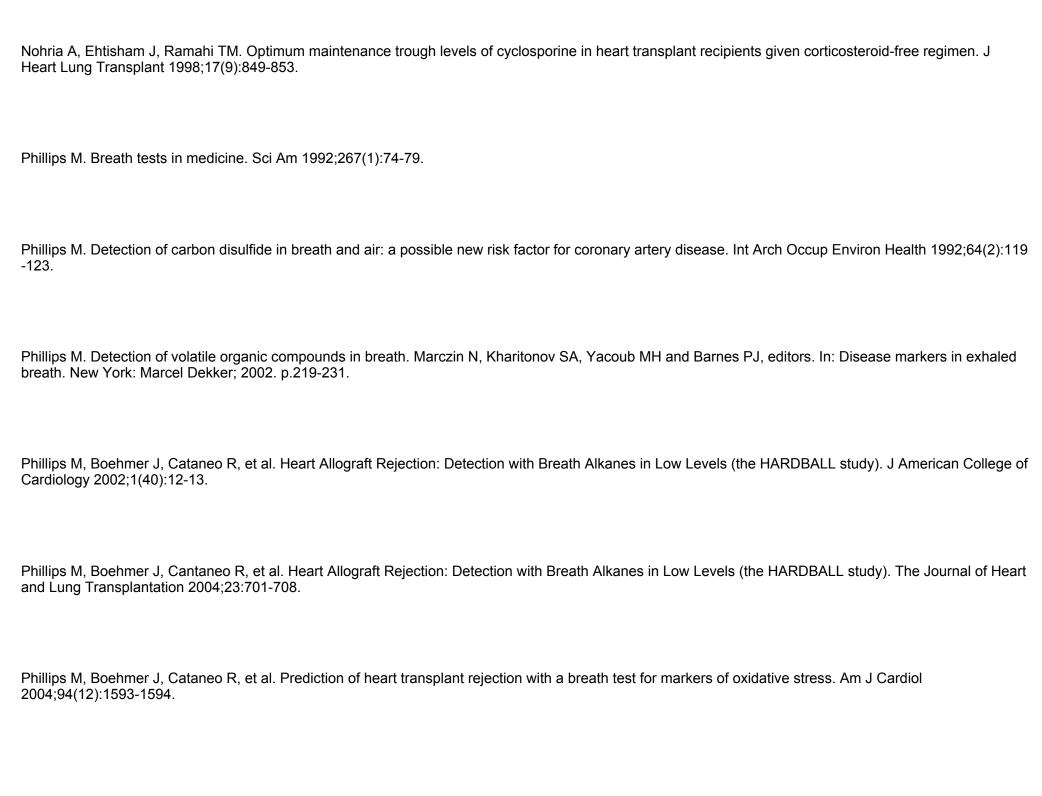


el Gamel A, Keevil B, Rahman A, Campbell C, Deiraniya A, Yonan N. Cardiac allograft rejection: do trough cyclosporine levels correlate with the grade of histologic rejection? J Heart Lung Transplant 1997;16(3):268-274.
Food and Drug Administration Heartsbreath test for heart transplant rejection. Approval Letter (Available at: http://www.fda.gov/cdrh/pdf3/H030004a.pdf).
Food and Drug Administration Heartsbreath test for heart transplant rejection. Center for Devices and Radiological Health: Consumer Information (Available at: http://www.fda.gov/cdrh/MDA/DOCS/H030004.html).
Food and Drug Administration. Heartsbreath test for heart transplant rejection. Summary of safety and probable benefit. (Available at: http://www.fda.gov/cdrh/pdf3/H030004b.pdf).
Food and Drug Administration The Heartsbreath Test: Product Labeling for Physicians (Available at: http://www.fda.gov/cdrh/pdf3/H030004c.pdf).
Food and Drug Administration Patient Information about the Heartsbreath Test (Available at: http://www.fda.gov/cdrh/pdf3/H030004d.pdf).
Food and Drug Administration. Humanitarian use devices. (Available at: http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfhde/HDEInformation.cfm).

Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Med Decis Making 1991;11(2):88-94.
Gvozdjakova A, Kucharska J, Mizera S, et al. Coenzyme Q10 depletion and mitochondrial energy disturbances in rejection development in patients after heart transplantation. Biofactors 1999;9(2-4):301-306.
Hausen B, Rhode R, Demertzis S, Albes JM, Wahlers T, Schafers HJ. Strategies for routine biopsies in heart transplantation based on 8-year results with more than 13,000 biopsies. Eur J Cardiothorac Surg 1995;9(10):592-598.
Hennekens CH, Buring JE. Mayrent SL, editor. Epidemiology in Medicine. Boston: Little, Brown and Company; 1987. 331-342.
Henzlova MJ, Nath H, Bucy RP, Bourge RC, Kirklin JK, Rogers WJ. Coronary artery to right ventricle fistula in heart transplant recipients: a complication of endomyocardial biopsy. J Am Coll Cardiol 1989;14(1):258-261.
Holt DW, Johnston A, Ramsey JD. Breath pentane and heart rejection. J Heart Lung Transplant 1994;13(6):1147-1148.
Chaw BA, Narula J. Antibody imaging in the evaluation of cardiovascular diseases. J Nucl Cardiol 1994; 1(5):457-476.
Kneepkens CM, Ferreira C, Lepage G, Roy CC. The hydrocarbon breath test in the study of lipid peroxidation: principles and practice. Clin Invest Med 1992;15(2):163-186.

Printed on 8/12/2011. Page 31 of 36

Kneepkens CM, Lepage G., Roy CC. The potential of the hydrocarbon breath test as a measure of lipid peroxidation. Free Radic Biol Med 1994;17(2):127-160.
Kohlmuller D, Kochen W. Is n-pentane really an index of lipid peroxidation in humans and animals? A methodological reevaluation. Anal Biochem 1993;210(2):268-276.
McAllister HA. Endomyocardial biopsy in the evaluation of conditions leading to cardiac transplantation and in the evaluation of cardiac allograft rejection. Texas Heart Institute Journal 1995;22(1):53-66.
Medicare Part B Extract and Summary System (BESS) Procedure Summary File for Heart Transplants in 2005, 2006 and 2007. Data pull on September 19 2008.
Mol BW, Lijmer JG, Evers JL, Bossuyt PM. Characteristics of good diagnostic studies. Semin Reprod Med 2003;21(1):17-25.
Moran AM, Lipshultz SE, Rifai N, O'Brien P, Mooney H, Perry S, et al. Non-invasive assessment of rejection in pediatric transplant patients: serologic and echocardiographic prediction of biopsy-proven myocardial rejection. J Heart Lung Transplant 2000;19(8):756-764.
Nielsen H, Sorensen FB, Nielsen B, Bagger JP, Thayssen P, Baandrup U. Reproducibility of the acute rejection diagnosis in human cardiac allografts. The Stanford Classification and the International Grading System. J Heart Lung Transplant 1993;12(2):239-243.



Phillips M, Greenberg J, Cataneo RN. Effect of age on the profile of alkanes in normal human breath. Free Radic Res 2000;33(1):57-63.
Phillips M, Cataneo R, Greenberg J, Grodman R, Salazar M. Breath markers of oxidative stress in patients with unstable angina. Heart Disease 2003;5(2):95-99.
Phillips M, Cataneo RN, Greenberg J, Gunawardena R, Naidu A, Rahbari-Oskoui F. Effect of age on the breath methylated alkane contour, a display of apparent new markers of oxidative stress. J Lab Clin Med 2000;136(3):243-249.
Phillips M, Cataneo R, Greenberg J, Gunawardena R, Rahbari-Oskoui F. Increased oxidative stress in younger as well as in older humans. Clinical Chimica Acta 2003;328: 83-86.
Phillips M, Cataneo R, Greenberg J, Grodman R, Gunawardena R, Naidu A. Effects of oxygen on breath markers of oxidative stress. European Respiratory Journal 2003;21(1):48-51.
Phillips M, Greenberg J, Sabas M. Alveolar gradient of pentane in normal human breath. Free Radic Res 1994;20(5):333-337.
Phillips M, Herrera J, Krishnan S, Zain M, Greenberg J, Cataneo RN. Variation in volatile organic compounds in the breath of normal humans. J Chronmatogr B Biomed Sci Appl 1999;729(1-2):75-88.
Schimke I, Schikora M, Meyer R, et al. Oxidative stress in the human heart is associated with changes in the antioxidative defense as shown after heart transplantation. Mol Cell Biochem 2000;204(1-2):89-96.
Printed on 8/12/2011. Page 34 of 36

Shanes JG, Ghali J, Billingham ME, et al. Interobserver variability in the pathologic interpretation of endomyocardial biopsy results. Circulation 1987;75(2):40-405.
Sobotka PA, Gupta DK, Lansky DM, Costanzo MR, Zarling EJ. Breath pentane is a marker of acute cardiac allograft rejection. J Heart Lung Transplant 1994;13(2):224-229.
Stewart S, Winters GL, Fishbein MC, et al. Revision of the 1990 working formulation for the standardization of nomenclature in the diagnosis of heart rejection J Heart Lung Transplant 2005;24(11):1710-1720.
Tanaka E. Invivo age-related changes in hepatic drug-oxidizing capacity in humans. J Clin Pharm Ther 1998;23(4):247-255.
Turgeon DK, Leichtman Ab, Lown KS, et al. P450 3A activity and cyclosporine dosing in kidney and heart transplant recipients. Clin Pharmacol Ther 1994;56(3):253-260.
Wagner K, Oliver MC, Boyle GJ, et al. Endomyocardial biopsy in pediatric heart transplant recipients: a useful exercise? (Analysis of 1,169 biopsies). Pediatr Transplant 2000;4(3):186-192.
Warnecke H, Muller J, Cohnert T, et al. Clinical heart transplantation without routine endomyocardial biopsy. J Heart Lung Transplant 1992;11(6):1093-1102.

Printed on 8/12/2011. Page 35 of 36

